Journal of Agricultural Machinery
Vol. 11, No. 2, Fall- Winter 2021, p. 371-383

نشریه ماشین‌های کشاورزی
جلد ۱۱، شماره ۲، نیمسال دوم ۱۴۰۰، ص ۳۸۳-۳۷۱

مقاله علمی-پژوهشی

انتخاب الگوریتم بهینه‌ی طبقه‌بندی و تشخیص تقلب در روغن زیتون با استفاده از ماشین‌بویایی

محمدرضا زارع زاده[۱]، محمد ابونجمی[۲*]، مهدی قاسمی ورنامخواستی[۳]، فاطمه آذری کیا[۴]

تاریخ دریافت: ۱۳۹۸/۰۸/۱۷
تاریخ پذیرش: ۱۳۹۸/۱۰/۲۵

چکیده

روغن زیتون فرابکر همیشه مورد توجه و خواست استفاده‌کنندگان می‌باشد؛ از این‌رو در روغن‌های زیتونِ بکر و فرابکر، تقلب، با افزودن روغن‌های با ارزش غذایی و قیمت کمتر مثل کانولا، آفتابگردان، تفاله زیتون و غیره مشاهده می‌گردد. در این پژوهش با استفاده از فناوری ماشین بویایی روغن زیتون فرابکر، از نمونه‌های تقلبی تهیه شده با روغن‌های متداول در بازار و با هفت مدل طبقه‌بندی مختلف تشخیص داده شد. نمونه‌ها در شش دسته‌ی "خالص و ۵، ۱۰، ۲۰، ۳۵ و ۵۰ درصد تقلب" و هر تیمار در هفت نمونه تهیه و آزمایش‌ها در هفت تکرار انجام گرفت. سامانه بویایی از هشت حسگر متفاوت تشکیل شده که برای هر کدام نمودار تغییر ولتاژ بر حسب زمان تهیه و از آن نمودار چهار ویژگی "کمینه، بیشینه، میانگین و اختلاف بیشینه و کمینه" در مدل‌های طبقه‌بندی استفاده گردید. به این ترتیب تعداد ۳۲ ویژگی برای هشت حسگر استخراج و مورد تجزیه و تحلیل قرار گرفت. به جهت تاثیر دمای نمونه‌ها در خروجی سامانه، تمام آزمایش‌ها در دمای ثابت انجام گرفتند. طبقه‌بندی نتایج با چهار الگوریتم طبقه‌بندی "K-همسایگی نزدیک، ماشین بردار پشتیبان، شبکه عصبی مصنوعی و آدابوست" صورت پذیرفت. ۷۰٪ داده‌ها برای آموزش و ۳۰٪ برای آزمون استفاده گردید. از بین ۳۲ ویژگی، ویژگی "کمینه مقدار خروجی سنسور TGS-822" بیشترین تاثیر را در دقت طبقه‌بندی داشتند. نتایج نشان داد روش طبقه‌بندی همسایگی نزدیک با بهترین دقت (۸۹/۸۹٪) و پس از آن روش ماشین بردار پشتیبان (۸۶/۵۲٪) بیشترین دقت طبقه‌بندی را دارا بودند.

واژه‌های کلیدی: تشخیص تقلب، روغن زیتون، طبقه‌بندی، کیفیت، ماشین بویایی

مقدمه

با گسترش پیشرفت‌های بشری در امر تغذیه و صنایع وابسته، امروزه امنیت و بهداشت تغذیه به‌عنوان یکی از اولویت‌های مهم جامعه بشری شناخته شده است (Borràs *et al.*, 2015). تقلب در مواد غذایی از جنبه‌های زیان اقتصادی به مصرف‌کننده، زیان آسیب به سلامتی و یا عدم بهره‌مندی از فواید تغذیه‌ای و جنبه‌ی روانی آن یعنی عدم اعتماد به محصول معرفی شده، نیازمند تحقیق و بررسی و نیز تصویب و اجرایی نمودن الزامات، قوانین و استانداردهای مربوط می‌باشند. تقلب در مواد غذایی مخصوصاً در روغن زیتون، با افزودن مواد با ارزش غذایی و قیمت پایین‌تر به روغن اصل ایجاد می‌گردد و این افزودن باعث تغییر در بعضی ویژگی‌ها می‌گردد. تشخیص روغن اضافه شده به روغن زیتون معمولاً به آسانی انجام نمی‌شود؛ چرا که گاهی روغن اضافه را طوری انتخاب می‌کنند که هیچ تغییری در شاخص‌ها (عدد یدی، عدد صابونی، تغییر ضریب شکست نور و غیره) به‌وجود نیاید (Ozen and Mauer, 2002; Pashaei and Shavakhi, 2017; Šmejkalová and Piccolo, 2010).

با توجه به این که استفاده از آزمون‌های بیولوژیکی آزمایشگاهی برای تشخیص آلاینده‌های غذایی امری زمان‌بر، پیچیده و در برخی موارد فاقد دقت کافی است، نیاز به فناوری‌های نوین ارزیابی برای حفاظت از زنجیره‌ی غذا بیش از پیش حس می‌شود (Samadi and Hosseyni nezhad, 2015). تحقیقات انجام یافته بر روی زیتون (*Olea europaea* L.) و روغن زیتون، در زمینه‌ی فواید تغذیه‌ای آن، باعث افزایش آگاهی و تمایل مردم نسبت به استفاده از آن شده و در نتیجه مقدار تولید آن در ایران و جهان افزایش قابل ملاحظه داشته است (Maghsudi, 1999). به‌طور کلی روغن زیتون به علت ثبات، ارزش غذایی، سازگاری، عطر، طعم و مزه‌ی آن به‌عنوان یک روغن با ارزش در نظر گرفته می‌شود. واردات سالانه‌ی روغن زیتون به ایران در مقیاس وسیع از کشورهای مدیترانه صورت می‌گیرد، هر چند مقداری از روغن‌های زیتون در ایران به‌ویژه در شمال تولید می‌شوند (Homapour *et al.*, 2016). در ایران مهم‌ترین استان‌های تولیدکننده‌ی زیتون به‌ترتیب گیلان، زنجان، کرمانشاه، مازندران و قزوین می‌باشند. به‌طور متوسط، هر دانه میوه‌ی زیتون حاوی ۲۰-۱۵٪ روغن می‌باشد. نوع و درصد ترکیبات شیمیایی موجود در روغن زیتون مانند ترکیب اسیدهای چرب، ترکیبات آنتی‌اکسیدانی و رنگدانه‌ها نماد کیفیت آن است که خود به چندین عامل مانند رقم،

۱- دانشجوی دکتری، گروه فنی کشاورزی، پردیس ابوریحان، دانشگاه تهران، ایران
۲- دانشیار، گروه فنی کشاورزی، پردیس ابوریحان، دانشگاه تهران، ایران
۳- دانشیار، گروه مهندسی مکانیک بیوسیستم، دانشگاه شهرکرد، شهرکرد، ایران
۴- استادیار، گروه فناوری صنایع غذایی، پردیس ابوریحان، دانشگاه تهران، ایران
(*- نویسنده مسئول: Email: abonajmi@ut.ac.ir)
DOI: 10.22067/jam.v11i2.84105

شرایط اقلیمی، روش استخراج و مرحله رسیدگی میوه زیتون بستگی دارد (Alavi Rafiee *et al.*, 2012).

مطابق تعاریف سازمان بین‌المللی روغن زیتون، IOOC[1] و همچنین سازمان ملی استاندارد ایران به شماره‌ی استاندارد ۱۴۴۶، روغن‌های زیتون به پنج دسته‌ی فرابکر (کمتر از ۰/۸٪ اسیدیته‌ی آزاد)، بکر (کمتر از ۲٪ اسیدیته‌ی آزاد)، بکر معمولی (کمتر از ۳/۳٪ اسیدیته‌ی آزاد)، بکر لامپاند (بیش از ۳/۳٪ اسیدیته‌ی آزاد) و پالایش شده (کمتر از ۰/۳٪ اسیدیته‌ی آزاد) طبقه‌بندی می‌شوند. اصلی‌ترین و بیشترین درصد اسیدچرب موجود در روغن زیتون اسید اولئیک نام دارد که از نام انگلیسی زیتون گرفته شده است. فرمول شیمیایی اسید اولئیک، $C_{18}H_{34}O_2$ می‌باشد و یک اسیدچرب تک غیراشباع می‌باشد.

تلاش‌های بسیاری جهت تشخیص تقلب در روغن‌های خوراکی صورت گرفته است. به‌عنوان مثال میلدنر و جلن جهت تشخیص تقلب در روغن زیتون مخلوط شده با روغن کلزا و آفتابگردان و استفاده کردند. (Mildner and Jelen, 2010). خو و همکاران نیز اکسید شدن انواع روغن‌های خوراکی را بررسی نمودند (Xu *et al.*, 2018). جهت تشخیص تقلب‌های به‌کار رفته در مواد غذایی و طبقه‌بندی آن‌ها، روش‌های مختلفی توسعه یافته‌اند که از جمله‌ی آن‌ها می‌توان به روش گرماسنجی (Fahim DaneshBahrami, 2015; Pashaei and Shavakhi, 2017)، کروماتوگرافی گازی (Lee *et al.*, 1998; Lee *et al.*, 2001)، کروماتوگرافی مایع با کارآیی بالا (HPLC) (Aued-Pimentel *et al.*, 2006; Seo *et al.*, 2010)، حسگرهای دی‌الکتریک (Naderi-Boldaji *et al.*, 2019)، روش‌های طیف‌نگاری ماوراء بنفش (UV) (Ok, 2017; Torrecilla *et al.*, 2010)، فروسرخ نزدیک (NIR) (Tay *et al.*, 2002) و فروسرخ میانه (MIR) و فروسرخ میانه فوریه (FT-MIR) (Rizki *et al.*, 2016; Zhao *et al.*, 2015)، رامان (Baeten and Aparicio, 2000)، رزونانس مغناطیسی هسته (NMR) (Sega *et al.*, 2010)، بینی و زبان الکترونیک (Haddi *et al.*, 2013; Martín *et al.*, 1999) و ماشین بینایی (Jafari *et al.*, 2014) اشاره نمود. این روش‌ها اغلب زمان‌بر بوده و نیاز به تجهیزات گران‌قیمت و کاربر متخصص دارند.

کاربردهای تشخیصی و طبقه‌بندی مواد غذایی به روش ماشین بویایی، به دلیل سادگی عملکرد، غیرمخرب بودن، سرعت و غیره روزبه روز درحال گسترش است. حاجی نژاد و همکاران با استفاده از سامانه ماشین‌بویایی توانستند تقلب در عسل را تشخیص دهند. آن‌ها از روش‌های تحلیل تفکیک خطی و شبکه عصبی مصنوعی (پس‌انتشار خطا) جهت طبقه‌بندی استفاده نمودند (Hajinezhad *et al.*, 2016). فیاض و همکاران نیز توسط سامانه ماشین‌بویایی مبتنی بر حسگرهای نیمه‌هادی اکسید فلزی به تشخیص و جداسازی اسانس‌های لیمو پرداختند و توانستند با دقت بالایی عمل تشخیص و طبقه‌بندی را انجام دهند (Fayyaz *et al.*, 2019). سامانه ماشین بویایی روشی بسیار موثر برای نظارت غیرمخرب بر رسیدگی میوه‌ها می‌باشد. ثنایی فر و همکاران میزان رسیدگی موز را با استفاده از حسگرهای بویایی نیمه‌هادی اکسید فلزی بررسی نمودند. ایشان دریافتند این حسگرها در ترکیب با روش‌های آنالیز تشخیص الگو توانایی آشکارسازی تغییرات رد اثر رسیدگی موز را دارد (Sanaeifar *et al.*, 2014). همچنین فساد گوشت چرخ شده نیز با استفاده از سامانه ماشین بویایی و با موفقیت بررسی و ارزیابی شده است (Varidi *et al.*, 2018). همچنین استخروئیه و راشدی با روش KNN و به روش ماشین‌بویایی توانستند نان‌های با کیفیت پایین را شناسایی کنند (Estakhroueiyeh and Rashedi, 2015).

یک سامانه ماشین بویایی می‌تواند ترکیب بودار را با تخمینی از غلظت آن و یا تعیین برخی از خواص ذاتی آن تشخیص دهد. انواع حسگرهای گازی که در قالب سیستم تشخیص یک دستگاه بینی الکترونیکی مورد استفاده قرار می‌گیرند، باید قادر به پاسخگویی به مولکول‌های موجود در فاز گازی باشند. این مولکول‌ها عمدتاً ترکیبات ارگانیک فرار[2] با جرم‌های مولی متفاوتی هستند. حسگرهای گازی در صنایع مختلفی از قبیل شیمیایی و پتروشیمی، محیط زیست، پزشکی و دارویی، کشاورزی، غذایی و بسیاری دیگر از فرآیندها که کنترل و تحلیل گاز در آن‌ها الزامی است، کاربرد دارند (Korotčenkov, 2013; Pearce *et al.*, 2006). با توجه به اهمیت کیفیت و ایمنی و سالم بودن روغن‌های گیاهی تولید شده در کارخانجات و همچنین افزایش تقلب در روغن‌های موجود در بازار مصرف، ضرورت تشخیص تقلب به روشی سریع و غیرمخرب در این محصولات بیش از پیش احساس می‌شود. آزمایش‌ها و روش‌های موجود علاوه بر زمان‌بر و هزینه‌بر بودن، مخرب نیز هستند. تقلب‌سنجی در آزمایشگاه‌های صنایع غذایی نیاز به تجهیزات آزمایشگاهی هزینه‌بر زیادی دارد، همچنین وقت‌گیر و مخرب است. در این پژوهش از حسگرهای نیمه‌هادی اکسید فلزی استفاده گردید. این حسگرها دارای قیمت کم، عمر زیاد، حساسیت بالا، پاسخ کم به رطوبت، پوشش طیف وسیع بوی محصولات کشاورزی و مواد غذایی، پاسخ سریع و توان تفکیک‌پذیری خوب می‌باشند.

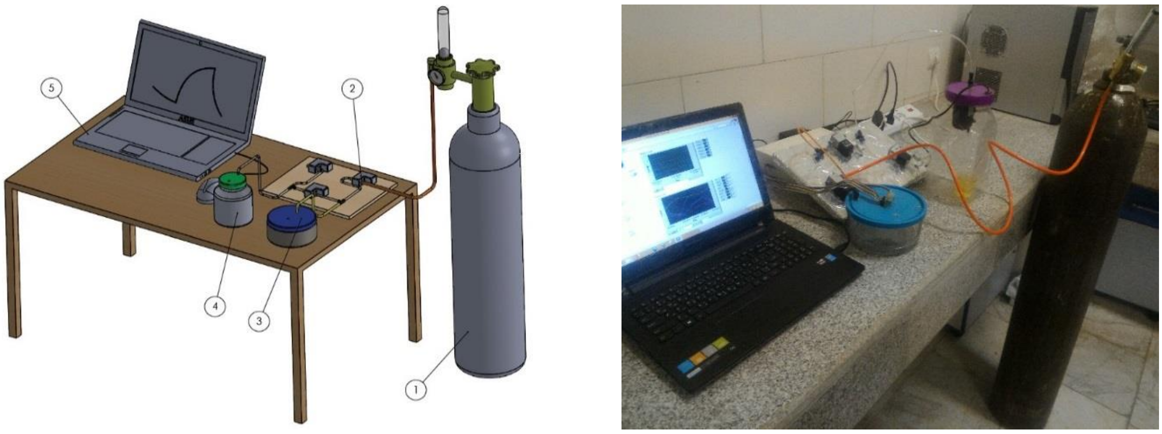
هدف از این پژوهش بررسی تقلب در روغن زیتون و انتخاب بهترین الگوریتم طبقه‌بندی و بهترین حسگر در ماشین بویایی می‌باشد. جهت انتخاب روش طبقه‌بندی بهینه، داده‌ها با چهار روش مختلف مدل شدند. همچنین از روش تحلیل مولفه‌های اصلی جهت حذف ویژگی‌های با درجه اهمیت کمتر استفاده گردید.

1- International Olive Oil Council

2- Volatile Organic Compounds

مواد و روش‌ها

در این پژوهش یک سامانه‌ی ماشین بویایی جهت دریافت مشخصات بویایی نمونه‌های مختلف روغن زیتون فرابکر و نمونه‌های تقلبی با درصدهای مختلف استفاده گردید. این مشخصات به منزله‌ی شناسامه و رد اثر هر نمونه بوده و منحصر به فرد می‌باشد (Ghasemi-Varnamkhasti *et al.*, 2015). شکل ۱ طرحواره‌ای از سامانه‌ی مورد استفاده را نشان می‌دهد. در این سامانه از کپسول اکسیژن (دارای یک فشارسنج در خروجی آن جهت نمایش فشار گاز خروجی)، شیرهای برقی، محفظه سنسورها، محفظه نمونه به همراه پمپ هوا در بالای محفظه، کارت تحصیل داده و مدارات الکتریکی و رایانه جهت مشاهده و ثبت داده‌ها استفاده می‌شود. آزمایش‌ها در گروه مکانیک بیوسیستم پردیس ابوریحان دانشگاه تهران، واقع در شهرستان پاکدشت انجام گرفتند.



الف ب

شکل ۱- سامانه‌ی ماشین بویایی مورد استفاده در تحقیق الف- تصویر واقعی، ب- طرحواره‌ی سامانه (۱- کپسول اکسیژن با فشارسنج، ۲- شیرهای برقی همراه دیتالاگر و مدارات الکتریکی مربوطه، ۳- محفظه حسگرها، ۴- محفظه قرارگیری نمونه‌ها همراه یک پمپ هوا، ۵- رایانه)

Fig.1. Used olfaction machine system a) Natural figure b) Schematic of system (1. Oxygen capsule 2. Solenoid valve, Data logger and electrical circuits 3. Sensors chamber 4. Samples chamber 5. PC)

تهیه نمونه‌ها

روغن زیتون فرابکر مورد استفاده در این پژوهش، از کارخانه صنایع غذایی دانزه در شهرستان لوشان استان گیلان تهیه گردید. ابتدا روغن‌کشی به‌صورت پرس سرد صورت گرفته و سپس جهت اطمینان از بکر بودن، مشخصات و درصدهای ترکیبات سازنده آن با سامانه کروماتوگرافی گازی (GC) مدل Agilent Technologies (7890B) در آزمایشگاه رهپویان دانش کولاک واقع در شهر تهران استخراج گردید. کروماتوگرافی گازی، روشی برای بررسی و جداسازی مواد فرار یک ماده که معمولاً به‌صورت مایع یا گاز است، می‌باشد. روغن زیتون فرابکر با روغن سرخ کردنی موجود در بازار که مخلوطی از روغن آفتاب‌گردان، کانولا (کلزا) و ذرت بود، با درصدهای وزنی ۵ و ۱۰ و ۲۰ و ۳۵ و ۵۰ درصد مخلوط شدند (شکل ۲).



شکل ۲- نمونه‌های روغن زیتون

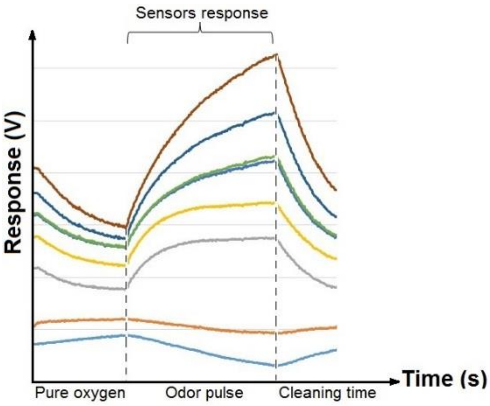
Fig.2. Olive oil samples

از هر دسته از تقلب‌ها، تعداد هفت نمونه‌ی ۱۰۰ گرمی تهیه گردیده و داده‌برداری در هفت تکرار صورت گرفت. به این ترتیب برای هر دسته تعداد ۴۹ آزمایش انجام پذیرفت. از عوامل مهم افت کیفیت روغن زیتون، نور، دما و هوا می‌باشد. بنابراین نمونه‌ها در محیط تاریک و در دمای کمتر از ۲۵ درجه سلسیوس نگهداری شدند و برای جلوگیری از ارتباط نمونه‌ها با هوا برای به حداقل رساندن سرعت اکسید شدن، ظرف‌های نگهداری نمونه‌های روغن زیتون، کاملاً پر تهیه گردید تا هوا در داخل ظرف و در مجاورت روغن زیتون کمینه گردد.

سامانه ماشین بویایی

روغن زیتون دارای ۲۶ نوع ترکیب فرار در چهار گروه آلدهیدها[1]، کتون‌ها[2]، الکل‌ها[3] و استرها[4] می‌باشد (Garrido-Delgado *et al.*, 2015). سامانه بینی الکترونیکی دارای یک محفظه‌ی نمونه با حجم کافی جهت قرارگیری نمونه‌ی روغن زیتون و جمع شدن مواد فرار ساطع شده از نمونه‌ها می‌باشد. در ابتدا نمونه، زمان معینی (۲۰ دقیقه) در داخل محفظه‌ی نمونه محبوس می‌شود. این عمل با هدف اشباع شدن محفظه از مواد فرار صورت گرفت. پس از زمان معین، مراحل کاری سامانه بویایی آغاز می‌شود. این مراحل شامل تصحیح خط مبنا، تزریق بوی نمونه به محفظه حسگرها و پاک‌سازی محفظه‌ی حسگرها و محفظه‌ی نمونه است. تغییرات در خروجی با تغییر ولتاژ خود را نشان می‌دهند. شکل ۳ نمونه‌ای از واکنش حسگرهای مورد استفاده در مراحل مختلف را نشان می‌دهد.

در این شکل محور افقی زمان و محور عمودی تغییر ولتاژ سنسور را نشان می‌دهد. زمان‌بندی برای تصحیح خط مبنا، زمان گذر بوی مد نظر از روی سنسورها و زمان پاک‌سازی سنسورها کاملاً تجربی می‌باشد. در این کار پس از ۲۰ دقیقه که نمونه داخل محفظه نمونه قرار گرفته و محفظه تقریباً از بوی روغن زیتون اشباع شده، ۱۵۰ ثانیه برای پاک‌سازی حسگرها با گاز اکسیژن، ۲۵۰ ثانیه زمان نمونه‌گیری یا همان زمان گذر ترکیبات فرار از روی حسگرها (انتقال هوای اشباع شده از ترکیبات فرار محفظه نمونه توسط پمپ هوای HAILEA مدل ACO-5501 با دبی ۱/۳ لیتر بر دقیقه) و ۱۰۰ ثانیه برای تصحیح خط مبنا در نظر گرفته شد. برای هر آزمایش، به مدت زمان هر یک ثانیه، داده‌برداری از خروجی‌های سنسورها صورت گرفت. فشار گاز عبوری از روی سنسورها که از یک کپسول اکسیژن تغذیه می‌نموده، با استفاده از فشارسنج تعبیه شده در خروجی آن، حدود ۲ بار در نظر گرفته شد. از آن‌جایی که حسگرهای مورد استفاده از نوع نیمه‌هادی بوده و این حسگرها در دمای بالاتر از دمای اتاق کارایی بهتری دارند، قبل از داده‌برداری، حسگرها پیش‌گرم شدند. سامانه ماشین بویایی مورد استفاده در این تحقیق شامل هشت نوع سنسور شامل دو سنسور گاز تاگوچی (TGS813 و TGS822 از شرکت فیگارو[5]، ژاپن)، ۵ سنسور گاز MQ (MQ3، MQ4، MQ8، MQ135، MQ136) از شرکت هانوی[6] چین و یک عدد سنسور FIS (شرکت نیسشا[7]، ژاپن) بود (شکل ۴). هرکدام از این حسگرها به طیفی از گازها (در این کار ترکیبات آروماتیک روغن زیتون) واکنش نشان می‌دهد.



شکل ۳- تغییرات خروجی حسگرها در مراحل مختلف برای نمونه EVOO

Fig.3. Sensor's output changes for EVOO



شکل ٤- حسگرهای مورد استفاده

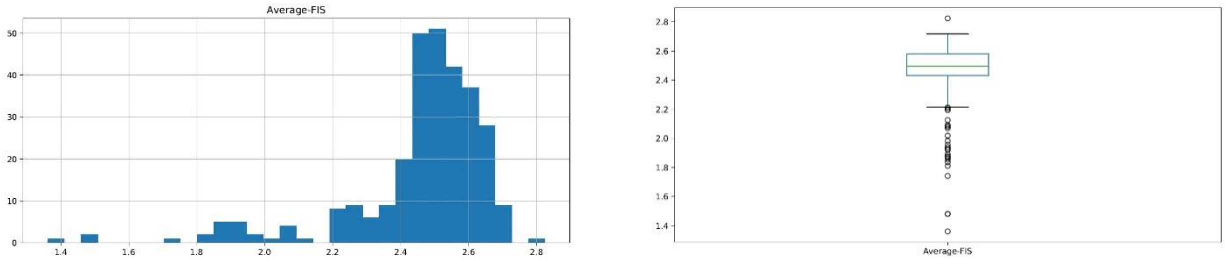
Fig.4. Used sensors

1- Aldehydes
2- Ketones
3- Alcohols
4- Esters
5- FIGARO
6- HANWEI
7- NISSHA

برای هر آزمون، چهار ویژگی برای هر حسگر استخراج گردید: ۱- خروجی حسگر در نقطه‌ی شروع مرحله‌ی نمونه‌گیری؛ ۲- خروجی حسگر در پایان مرحله‌ی نمونه‌گیری؛ ۳- تفاضل خروجی حسگر در ابتدا و انتهای نمونه‌گیری و ۴- میانگین خروجی حسگر در مرحله‌ی نمونه‌گیری به این ترتیب با داشتن هشت عدد سنسور، خروجی در هر مرحله از داده‌برداری شامل هشت منحنی از منحنی‌های مشابه نمودار شکل ۳ خواهد بود و بنابراین برای هر مرحله آزمون، تعداد ۳۲ ویژگی بویایی استخراج گردید. خروجی‌های هشت حسگرِ مورد اشاره، توسط کانال‌های آنالوگ کارت تحصیل داده و انتقال آن از طریق پورت USB به رایانه و محیط نرم‌افزار LabView وارد شدند. اطلاعات در رایانه به‌صورت فایل اکسل ذخیره شدند. همچنین باز و بسته نمودن شیرهای برقی و عملکرد پمپ و زمان‌بندی آن‌ها توسط برنامه نوشته شده با نرم‌افزار LabView و به‌وسیله‌ی کانال‌های دیجیتال کارت تحصیل داده انجام گرفت.

تحلیل‌های آماری

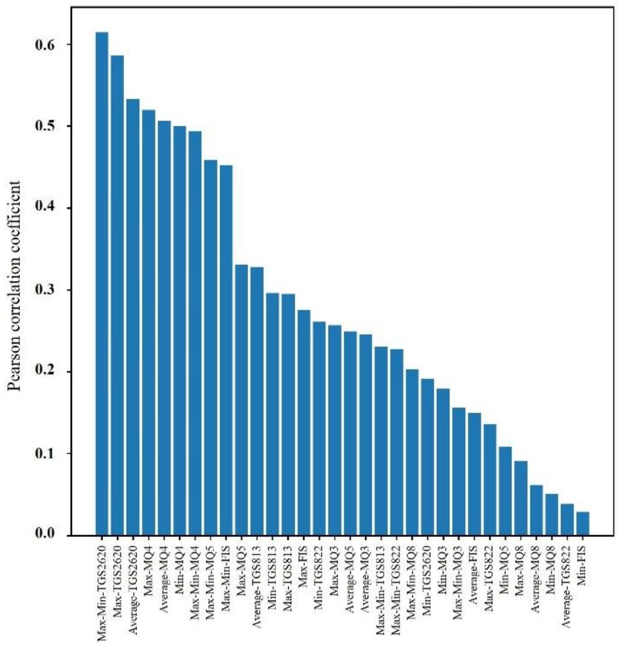
برای اطلاع از توزیع و پراکندگی داده‌ها، تمام داده‌های مربوط به تک‌تک ۳۲ ویژگی بویایی، با نمودارهای جعبه‌ای و فراوانی مطابق با نمودارهای شکل ۵ واکاوی شدند.



ب الف

شکل ۵- نمودارهای پراکندگی داده‌ها برای یک داده نمونه الف) نمودار جعبه‌ای، ب) نمودار فراوانی

Fig.5. Distribution diagram for a sample feature a) Box plot, b) Histogram



شکل ٦- همبستگی ویژگی‌ها با دسته

Fig.6. Features correlation with type

در نمودار باکس پلات ۵-الف مشاهده می‌شود که داده‌ها هم از بالا و هم از پایین، تعدادی داده دورافتاده یا پرت[1] می‌باشند. همچنین جعبه که چارک اول تا سوم را پوشش می‌دهد بخش کوچکی را شامل می‌شود و لذا تمرکز و فراوانی داده‌ها در یک بازه‌ی خاصی اتفاق افتاده است. این موارد نشان می‌دهد داده‌ها توزیع مناسب و نرمالی

1- Outlier

نداشته و نیاز به عملیات پیش‌پردازش خواهند داشت. این اطلاعات از نمودار ۵-ب نیز به وضوح قابل استخراج و دریافت می‌باشد.

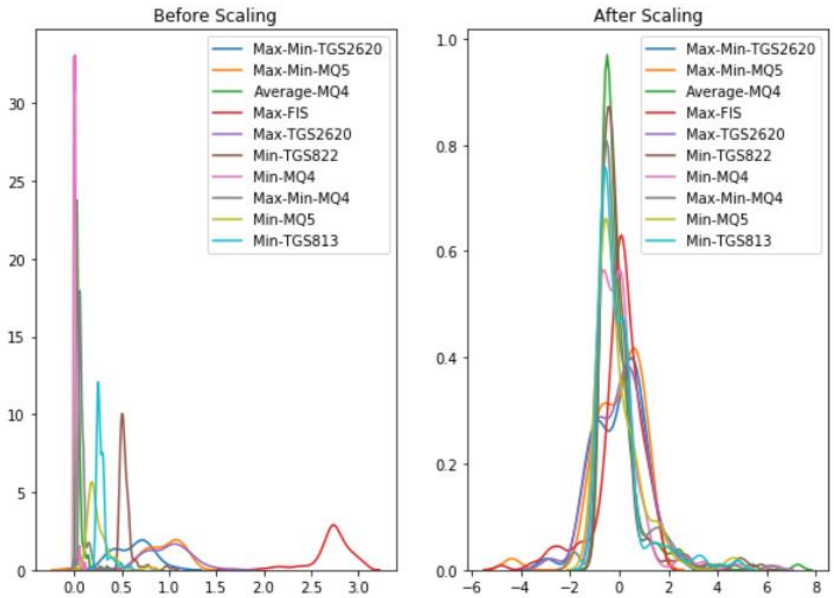
در این تحقیق، ارتباط ویژگی‌های بویایی با طبقه‌ی مورد نظر، (ضریب همبستگی) مقایسه گردید (شکل ۶). بر اساس این نمودار، ویژگیِ "اختلاف بیشینه و کمینه مقدار خروجی حسگر TGS-2620" بیشترین همبستگی را با طبقه مورد نظر نشان داد.

پیش‌پردازش

برای استفاده موثرتر از داده‌های خام، پیش‌پردازش‌های لازم روی داده‌ها صورت گرفت. پیش‌پردازش، عملیات داده‌کاوی اولیه می‌باشد و با این کار ویژگی‌های مستخرج از سامانه‌ها به‌صورت مفیدتر استفاده می‌گردد (Pyle, 1999). نرمال کردن یکی از مهم‌ترین عملیات پیش‌پردازش می‌باشد. متغیر نرمال استاندارد، یک متغیر طبیعی با میانگین برابر صفر (μ=0) و انحراف استاندارد برابر ۱ (σ=1) با یک تابع احتمال است. در این کار چون داده‌ها مطابق بررسی‌های اولیه (نمودارهای جعبه‌ای و هیستوگرام در شکل ۵) نرمال نبودند، قبل از کار مدل‌سازی و طبقه‌بندی داده‌ها، داده‌ها با استفاده از رابطه (۱)، استاندارد شدند.

$$z=\frac{(x-\mu)}{s} \qquad (۱)$$

که در آن z داده‌ی استاندارد شده، x داده‌ی آموزش قبل از استاندارد سازی، μ میانگین داده‌ها و s انحراف معیار داده‌های آموزش می‌باشند. شکل ۷ به‌صورت شماتیک نشان می‌دهد که استانداردسازی چه اثری بر وضعیت داده‌ها می‌گذارد.



شکل ۷- شماتیک نرمال‌سازی ده ویژگی مهم استخراج شده

Fig.7. Schematic of standardization of first 10 features

زمانی‌که ویژگی‌ها و داده‌ها با مقیاس‌های مختلف می‌باشند، استانداردسازی بسیار حائز اهمیت می‌باشند.

در این پژوهش، کدهای پیش‌پردازش داده‌ها و طبقه‌بندی داده‌ها با استفاده از نرم‌افزار پایتون[1] نسخه ۳.۷ نوشته و اجرا شدند.

نتایج و بحث

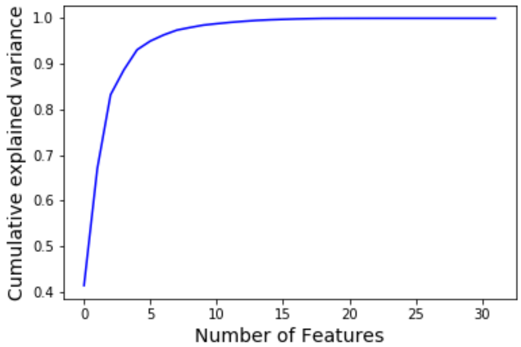
استخراج ویژگی‌های موثر

به‌دلیل تعداد زیاد ویژگی‌های این پژوهش (۳۲ ویژگی)، برای کاستن ویژگی‌ها، باید با روش‌های آماری، مهم‌ترین ویژگی‌ها را تعیین و بر آن‌ها تمرکز نمود. لذا در این مقاله روش تحلیل مولفه‌های اصلی[2] (PCA)، جهت تعیین مهم‌ترین ویژگی‌های موثر در طبقه بندی و کاهش ابعاد داده‌ها به‌کار گرفته شد (شکل ۸).

همان‌طور که از شکل ۸ (نتیجه‌ی تحلیل PCA) بر می‌آید، با در دست داشتن حدود پنج ویژگی، حدود بیش از ۹۵٪ دقت به‌دست خواهد آمد. مشاهده می‌شود که با افزایش تعداد ویژگی‌ها از حدود ۱۰-۱۲ عدد، شیب نمودار تغییر محسوسی نداشته و بنابراین تعداد کمتر ویژگی‌ها نیز ما را به هدف خواهد رساند. در ادامه، ویژگی‌های برتر مورد استفاده (ویژگی دارای بیشترین همبستگی با دسته‌ی مورد نظر) مطابق شکل ۹ نشان داده شده‌اند.

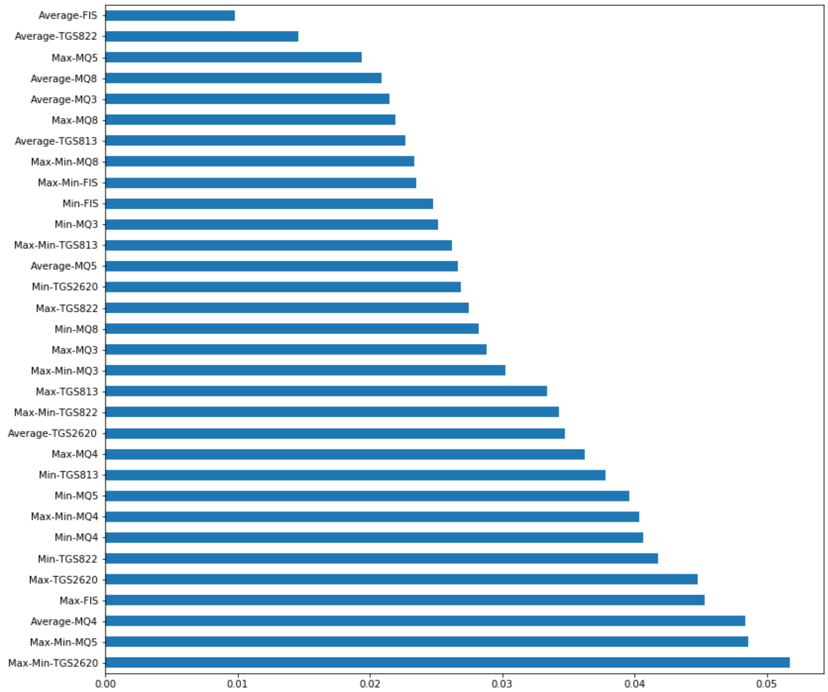
1- Python

2- Principal Component Analysis



شکل ۸- نتیجه تحلیل PCA برای استخراج مهم‌ترین ویژگی‌ها

Fig.8. PCA result for important features extracting



شکل ۹- لیست ویژگی‌ها به‌ترتیب اهمیت با الگوریتم‌های مبتنی بر درخت

Fig.9. Features list sorted by importance with extra trees classifier

برای استخراج ترتیب اهمیت ویژگی‌ها از شیوه‌ی طبقه‌بندی کننده‌ی درخت‌های شدیداً تصادفی شده[1] استفاده گردید. جدول ۱، مهم‌ترین ویژگی‌ها (ویژگی‌های با بیشترین همبستگی با دسته) را بر حسب درصد نشان می‌دهد.

جدول ۱ نشان می‌دهد ویژگی "اختلاف بیشینه و کمینه‌ی حسگر TGS2620" و "اختلاف بیشینه و کمینه حسگر MQ5" و "میانگین خروجی‌های حسگر MQ4" به‌ترتیب سه ویژگی دارای تاثیر بالا بر روی دقت طبقه‌بندی می‌باشند، لذا سه حسگر TGS2620، MQ5 و MQ4 حسگرهای تاثیرگذار و مهم در این آزمون می‌باشند. در جمع‌بندی کلی و از آنجا که نتایج تحلیل مولفه‌های اصلی نشان‌دهنده تاثیر بیشتر حسگر TGS2620 در تشخیص تقلب و طبقه‌بندی روغن‌های زیتون بود و چون کاربرد اصلی این حسگر در الکل، تولوئن و دیگر بخارات آلی و الکل در محدوده‌ی غلظت ۵۰ تا ۵۰۰۰ ppm می‌باشد، می‌توان نتیجه گرفت احتمالاً این ترکیبات آروماتیک به‌عنوان شناساگر روغن‌های تقلبی از فرابکر موثرتر عمل می‌کنند و تغییرات آن‌ها به درصد، با افزایش درصد تقلب در روغن زیتون فرابکر قابل توجه بوده و البته در همان محدوده‌ی 50-5000ppm می‌باشند. همچنین حسگرهایی مثل MQ3 و MQ8 تاثیر به مراتب کمتری در طبقه‌بندی داشتند. کاربرد اصلی سنسور MQ8 برای هیدروژن و نیز الکل می‌باشد. همچنین سنسور MQ3 نیز در محدوده‌ی 0.05-10 ppm برای الکل واکنش نشان می‌دهد و چون این دو سنسور تاثیر کمتری در طبقه‌بندی داشتند، لذا نتیجه گرفته می‌شود که گازهای

1- Extra Tree Classifier (ETC)

نامبرده، در بین نمونه‌های این تحقیق تغییر محسوسی در دامنه مورد اشاره نداشتند.

جدول ۱- ویژگی‌های مهم در طبقه‌بندی

Table 1- Important features in classification

Feature	Importance (٪)
Max-Min (TGS2620)	5.17
Max-Min (MQ5)	4.86
Average (MQ4)	4.84
Max (FIS)	4.53
Max (TGS2620)	4.48
Min (TGS822)	4.18
Min (MQ4)	4.07
Max-Min (MQ4)	4.03
Min (MQ5)	3.96
Min (TGS813)	3.78

طبقه‌بندی

جهت اعتبارسنجی مدل‌های طبقه‌بندی، بایستی داده‌ها به دو دسته‌ی آموزش و آزمون تقسیم شوند. در این کار ما ۷۰ درصد داده‌ها را برای آموزش و ۳۰ درصد را برای آزمون استفاده کردیم.

روش K-همسایگی نزدیک

روش طبقه‌بندی K-همسایگی نزدیک با ۸۹/۸۹٪ دقت در طبقه‌بندی داده‌های آزمون و ۹۴/۶۳٪ دقت در طبقه‌بندی داده‌های آموزش، بهترین مدل طبقه‌بندی در این آزمون می‌باشند. روش K-همسایگی نزدیک یک روش یادگیری موردی است و از ساده‌ترین روش‌های طبقه‌بندی یادگیری ماشین به حساب می‌آید. در این روش، یک نمونه با رای اکثریت از همسایه‌هایش دسته‌بندی می‌شود. روش K-همسایگی نزدیک، برای بسیاری از روش‌ها کاربرد دارد، زیرا اثربخش، غیرپارامتریک و دارای پیاده‌سازی راحت می‌باشد. بهترین انتخاب از K، وابسته به داده‌ها می‌باشد؛ به‌طور کلی مقدار بزرگ از K، اثر نویز روی دسته‌بندی را کاهش می‌دهد، اما مرز مابین کلاس‌ها کمتر متمایز می‌شود. Leaf size در این روش ۳۰ در نظر گرفته شد. این روش طبقه‌بندی دقت بالایی را ارائه نمود که به دلیل زیاد بودن تعداد ویژگی‌ها، این نتیجه قابل پیش‌بینی بود.

جدول ۲- ماتریس اغتشاش برای روش KNN

Table 2- Confusion matrix of KNN

Classes	0%	5%	10%	20%	35%	50%
0%	17	0	0	0	0	0
5%	2	16	0	0	0	0
10%	0	0	10	2	0	1
20%	0	0	0	11	1	0
35%	0	0	2	0	8	1
50%	0	0	0	0	0	16

ماتریس اغتشاش تشخیص روغن زیتون تقلبی برای روش KNN در جدول ۲ آورده شده است که حاکی از این است که این روش قادر به تشخیص نمونه‌های تقلبی با دقت بالا می‌باشد.

روش ماشین بردار پشتیبان

پس از روش KNN، مدل طبقه‌بندیِ SVM (با ۸۶/۵۲٪) به‌عنوان بیشترین دقت طبقه‌بندی تعیین شد. وجنوسکی و همکاران و نیز خو و همکاران با همین روش طبقه‌بندی، روغن زیتون را از لحاظ کیفی با دقت بالایی نسبت به سایر روش‌ها طبقه‌بندی نمودند (Wojnowski *et al.*, 2017; Xu *el al.*, 2018). همچنین نادری بلداجی و همکاران نیز از روش ماشین بردار پشتیبان و نیز روش تحلیل مولفه‌های اصلی (PCA) توانستند تقلب در نمونه‌های روغن کنجد را به‌دست آورند (Naderi-Boldaji *el al.*, 2019). ماشین بردار پشتیبان (SVM) روشی برای طبقه‌بندی و مسائل رگرسیون می‌باشد. در طبقه‌بندی، SVM بر به‌دست آوردن مرز "بهینه" دو کلاس در فضای بردار در مجموعه داده‌ها متمرکز است. می‌توان گفت این روش، از مرسوم‌ترین روش‌های طبقه‌بندی در کیفیت‌سنجی صنایع غذایی می‌باشد.

جدول ۳- ماتریس اغتشاش برای روش SVM

Table 3- Confusion matrix of SVM

Classes	0%	5%	10%	20%	35%	50%
0%	17	0	0	0	0	0
5%	1	15	2	0	0	0
10%	0	0	12	1	2	0
20%	0	0	0	11	1	0
35%	0	0	0	0	10	1
50%	0	0	0	0	1	15

ماتریس اغتشاش تشخیص روغن زیتون تقلبی برای روش SVM در جدول ۳ آورده شده است. مشاهده می‌شود که دقت طبقه‌بندی این روش نیز همانند روش قبلی مناسب و دقیق می‌باشد. از نتایج تحقیقات مختلف در طبقه‌بندی انواع روغن استنتاج می‌شود که روش‌هایی مثل SVM و KNN با داشتن دقت بالا، روش توصیه شده‌تری برای این گونه تحقیقات باشد.

روش شبکه عصبی مصنوعی

در این کار از شبکه عصبی مصنوعی MLP دارای یک لایه مخفی (لایه میانی) با یک‌صد نرون استفاده گردید. روش طبقه‌بندی شبکه عصبی مصنوعی با ۶۷/۴۲ درصد، جزء مدل‌های کم دقت به‌دست آمد. احتمالاً در صورت افزایش تعداد داده‌ها برای هر دسته، دقت بالاتری به‌دست خواهد آمد، زیرا روش شبکه‌ی عصبی مصنوعی با افزایش تعداد داده، نتیجه‌ی بهتری ارائه می‌دهد. جهت تشخیص تقلب در روغن زیتون فرابکر، خو و همکاران با سامانه‌ی طیف‌سنجی تبدیل فوریه مادون قرمز (FTIR Spectroscopy) و از روش‌های LDA، شبکه عصبی مصنوعی و ماشین بردار پشتیبان استفاده نمودند. در تحقیق آن‌ها نیز روش شبکه عصبی مصنوعی، کمترین دقت طبقه‌بندی را (همانند نتایج این پژوهش) داشت (Xu *et al.*, 2018). همچنین اولیورز و همکاران نیز به پایین بودن دقت طبقه‌بندی به روش شبکه عصبی مصنوعی اذعان کردند. آن‌ها از روش‌های LDA، QDA و ANN برای طبقه‌بندی داده‌های سامانه‌ی ماشین‌بویایی جهت تشخیص تقلب در روغن زیتون استفاده نمودند (Oliveros *et al.*, 2002). البته نتایج مشابهی را دانکوفسکا و کاوالوسکی در بررسی تازگی و نیز تشخیص نوع روغن زیتون به روش طیف‌های فلورسانس گرفتند. آن‌ها روش‌های مختلفی را برای طبقه‌بندی به‌کار بردند که از بین آن‌ها روش‌های KNN و SVM (مشابه نتایج این پژوهش) دارای بالاترین دقت بودند (Dankowska and Kowalewski, 2019). می‌توان نتیجه گرفت احتمالاً روش شبکه عصبی مصنوعی در کاربردهای تشخیص تقلب و طبقه‌بندی کیفی مواد غذایی تاثیر کمتری (نسبت به سایر روش‌های طبقه‌بندی) ارائه می‌نماید. البته در بعضی پژوهش‌ها روش شبکه عصبی از بهترین و دقیق‌ترین روش‌های طبقه‌بندی به‌دست آمده است (Hajinezhad *et al.*, 2016).

جدول ٤- ماتریس اغتشاش برای روش ANN

Table 4- Confusion matrix of ANN

Classes	0%	5%	10%	20%	35%	50%
0%	17	0	0	0	0	0
5%	13	2	2	1	0	0
10%	3	0	7	1	4	0
20%	0	0	0	11	1	0
35%	0	0	0	1	6	4
50%	0	0	0	0	1	14

روش طبقه‌بندی آدابوست

آدابوست یک الگوریتم یادگیری ماشین و در واقع الگوریتمی به‌منظور ارتقاء عملکرد و رفع مشکل رده‌های نامتوازن بوده که همراه دیگر الگوریتم‌های یادگیری استفاده می‌شود. آدابوست نسبت به داده‌های نویزی و پرت حساس است؛ ولی نسبت به مشکل بیش برازش از بیشتر الگوریتم‌های یادگیری برتری دارد. به دلیل مناسب نبودن پراکندگی و توزیع داده‌ها (که در تحلیل‌های اولیه مطابق شکل ۵ مشخص شده بود)، کمترین دقت طبقه‌بندی از بین چهار روش

مورد بررسی، به روش آدابوست مربوط است. روش آدابوست، یک روش کم‌کاربرد در طبقه‌بندی مواد غذایی می‌باشد. البته در مواردی مثل شناسایی نوع روغن خوراکی در این حوزه موثر واقع شده است (Huo *et al.*, 2012).

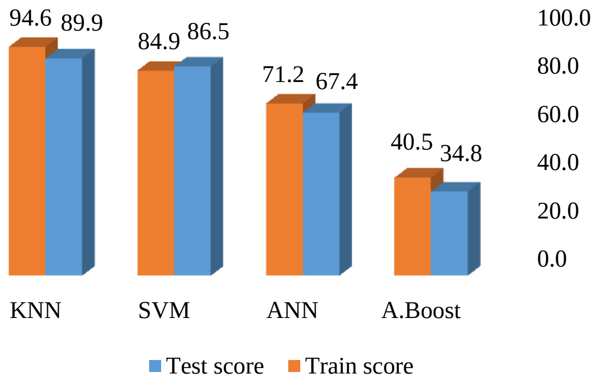
جدول ۵- ماتریس اغتشاش برای روش AdaBoost

Table 5- Confusion matrix of AdaBoost

Classes	0%	5%	10%	20%	35%	50%
0%	0	13	0	0	0	4
5%	0	2	0	0	0	16
10%	0	4	0	10	0	1
20%	0	0	0	12	0	0
35%	0	0	0	10	0	1
50%	0	0	0	15	0	1

از ماتریس اغتشاشِ این روش نیز به وضوح مناسب نبودن این روش در طبقه‌بندی استنتاج می‌شود.

نتایج طبقه‌بندی داده‌ها با مدل‌های اشاره شده در بالا در شکل ۱۰ آورده شده است. همان‌طور که از این جدول برمی‌آید، روش KNN و SVM بالاترین دقت طبقه‌بندی را ارائه داده‌اند.



شکل ۱۰- نتایج طبقه‌بندی با ۴ مدل مختلف

Fig.10. Classification results with 4 different models

نتیجه‌گیری

در این پژوهش یک سامانه‌ی ماشین بویایی ساده و کم‌هزینه بر اساس هشت حسگر نیمه‌هادی اکسید فلزی برای تشخیص تقلب و طبقه‌بندی روغن زیتون مورد استفاده قرار گرفت. این روش آزمون، غیرمخرب بوده و برخلاف برخی روش‌های ارزیابی مواد غذایی، به امکانات پیچیده و گران‌قیمت آزمایشگاهی نیاز ندارد. طبقه‌بندی خروجی‌های پیش‌پردازش شده، با چهار مدل دسته‌بندی "K-همسایگی نزدیک، ماشین بردار پشتیبان، شبکه عصبی مصنوعی و آدابوست" با استفاده از نرم‌افزار پایتون ۳.۷ صورت پذیرفت. از بین این مدل‌های طبقه‌بندی، مدل طبقه‌بندی K-همسایگی نزدیک با ۸۹/۸۹ درصد دارای بیشترین دقت بود. همچنین مدل‌های ماشین بردار پشتیبان و شبکه عصبی مصنوعی به‌ترتیب با ۸۶/۵۲٪ و ۶۷/۴۲٪ از لحاظ دقت، دومین و سومین مدل طبقه‌بندی تعیین شدند. نتایج بررسی و تحلیل داده‌های حاصل از سامانه بویایی نشان داد از بین هشت عدد سنسور بویایی، حسگرهای TGS2620، MQ5 و MQ4 سه حسگرِ با بالاترین تاثیر در طبقه‌بندی بودند. می‌توان نتیجه‌گیری کرد تقلب‌های روغن زیتون فرابکر با درصدهای پایین (پنج درصد) به روش ماشین‌بویایی قابل تشخیص می‌باشد.

References

1. Alavi Rafiee, S., R. Farhoosh, and M. Haddad Khodaparast. 2012. Physicochemical properties of Iranian commercial olive oils. Iranian Journal of Nutrition Sciences & Food Technology 7: 85-94. (In Farsi)

2. Aued-Pimentel, S., E. Takemoto, R. Antoniassi, and E. S. G. Badolato. 2006. Composition of tocopherols in sesame seed oil: an indicative of adulteration. Grasas y Aceites 57: 205-210.
3. Baeten, V., and R. Aparicio. 2000. Edible oils and fats authentication by Fourier transform Raman spectrometry. BASE.
4. Borràs, E., J. Ferré, R. Boqué, M. Mestres, L. Aceña, and O. Busto. 2015. Data fusion methodologies for food and beverage authentication and quality assessment–A review. Analytica Chimica Acta 891: 1-14.
5. Dankowska, A., and W. Kowalewski. 2019. Comparison of different classification methods for analyzing fluorescence spectra to characterize type and freshness of olive oils. European Food Research and Technology 245: 745-752.
6. Estakhroueiyeh, H. R., and E. Rashedi. 2015. Detecting moldy bread using an e-nose and the KNN classifier. Pages 251-255. 2015 5th International Conference on Computer and Knowledge Engineering (ICCKE): IEEE.
7. Fahim Danesh, M., and M. E. Bahrami. 2015. Evaluation of adulteration in sesame oil using Differential Scanning Calorimetry. Food Science and Technology 13: 81-89. (In Farsi)
8. Fayyaz, P., S. S. Mohtasebi, A. Jafari, and A. Masoudi. 2019. Development and Evaluation of an Electronic Nose System Based on MOS Sensors to Detect and to Distinguish Lemon Essential Oils. Journal of Agricultural Machinery 9. (In Farsi)
9. Garrido-Delgado, R., M. del Mar Dobao-Prieto, L. Arce, and M. Valcárcel. 2015. Determination of volatile compounds by GC–IMS to assign the quality of virgin olive oil. Food Chemistry 187: 572-579.
10. Ghasemi-Varnamkhasti, M., S. S. Mohtasebi, M. Siadat, H. Ahmadi, and S. H. Razavi. 2015. From simple classification methods to machine learning for the binary discrimination of beers using electronic nose data. Engineering in Agriculture, Environment and Food 8: 44-51.
11. Haddi, Z., H. Alami, N. El Bari, M. Tounsi, H. Barhoumi, A. Maaref, N. Jaffrezic-Renault, and B. Bouchikhi. 2013. Electronic nose and tongue combination for improved classification of Moroccan virgin olive oil profiles. Food Research International 54: 1488-1498.
12. Hajinezhad, M., S. S. Mohtasebi, M. Ghasemi-varnamkhasti, and M. Aghbashlo. 2016. Detecting Adulteration in Lotus Honey Using a Machine Olfactory System. Journal of Agricultural Machinery 7: 439-450. (In Farsi)
13. Homapour, M., M. Ghavami, Z. Piravivanak, and E. Hosseini. 2016. Evaluation of Chemical Characteristics of Extra Virgin Olive Oils Extracted from Three Monovarieties of Mari, Arbequina and Koroneiki in Fadak and Gilvan Regions. Journal of Food Biosciences and Technology 6: 77-85.
14. Huo, Q.-G., X.-B. Jin, and H.-M. Zhang. 2012. Multi-label classification for Oil Authentication. Pages 711-714. 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery: IEEE.
15. Jafari, A., A. Fazayeli, and M. R. Zarezadeh. 2014. Estimation of orange skin thickness based on visual texture coarseness. Biosystems Engineering 117: 73-82.
16. Korotčenkov, G. S. 2013. Handbook of Gas Sensor Materials: Properties, Advantages and Shortcomings for Applications. Conventional Approaches. Springer.
17. Lee, D.-S., B.-S. Noh, S.-Y. Bae, and K. Kim. 1998. Characterization of fatty acids composition in vegetable oils by gas chromatography and chemometrics. Analytica Chimica Acta 358: 163-175.
18. Lee, D.-S., E.-S. Lee, H.-J. Kim, S.-O. Kim, and K. Kim. 2001. Reversed phase liquid chromatographic determination of triacylglycerol composition in sesame oils and the chemometric detection of adulteration. Analytica Chimica Acta 429: 321-330.
19. Maghsudi, S. 1999. Technology of olive and it's products Tehran: Iran Agriculture Science Press. (In Farsi)
20. Martín, Y. G., J. L. P. Pavon, B. M. Cordero, and C. G. A. Pinto. 1999. Classification of vegetable oils by linear discriminant analysis of electronic nose data. Analytica Chimica Acta 384: 83-94.
21. Mildner‐Szkudlarz, S., and H. H. Jeleń. 2010. Detection of olive oil adulteration with rapeseed and sunflower oils using mos electronic nose and SMPE‐MS. Journal of Food Quality 33: 21-41.
22. Naderi-Boldaji, M., M. Mokhtari, M. Ghasemi-Varnamkhasti, and M. Tohidi. 2019. Feasibility of using a cylindrical resonator sensor for adulteration detection in sesame oil. Innovative Food Technologies (JIFT).
23. Ok, S. 2017. Detection of olive oil adulteration by low-field NMR relaxometry and UV-Vis spectroscopy upon mixing olive oil with various edible oils. Grasas y Aceites 68: 173.
24. Oliveros, M. C. C., J. L. P. Pavón, C. G. A. Pinto, M. E. F. Laespada, B. M. Cordero, and M. Forina. 2002. Electronic nose based on metal oxide semiconductor sensors as a fast alternative for the detection of adulteration of virgin olive oils. Analytica Chimica Acta 459: 219-228.
25. Ozen, B. F., and L. J. Mauer. 2002. Detection of hazelnut oil adulteration using FT-IR spectroscopy. Journal of Agricultural and Food Chemistry 50: 3898-3901.
26. Pashaei, A., and F. Shavakhi. 2017. Application of calorimeter for detection of adulteration of extra virgin olive oil mixed with aspersed oil as a substitute for gas chromatography. Journal of Food Research (Agricultural Science) 27: 139-149. (In Farsi)
27. Pearce, T. C., S. S. Schiffman, H. T. Nagle, and J. W. Gardner. 2006. Handbook of machine olfaction: electronic nose technology. John Wiley & Sons.

28. Pyle, D. 1999. Data preparation for data mining. morgan kaufmann.
29. Rizki, H., W. Terouzi, F. Kzaiber, H. Hanine, and A. Oussama. 2016. Quantification of Adulterations in Sesame Oil with Inferior Edible Oils by using ATR-FTIR Coupled to Chemometrics. IOSR Journal of Environmental Science, Toxicology and Food Technology, e-ISSN: 2319-2402.
30. Samadi, S., and M. Hosseyni nezhad. 2015. Smart sensors in the food industry: Technology and applications. (In Farsi)
31. Sanaeifar, A., S. Mohtasebi, M. Ghasemi-Varnamkhasti, and H. Ahmadi. 2014. Design, Construction and Performance Evaluation of a Metal Oxide Semiconductor (MOS) Based Machine Olfaction (Electronic Nose) for Monitoring of Banana Ripeness. Journal of Agricultural Machinery 5: 111-121.
32. Sega, A., I. Zanardi, L. Chiasserini, A. Gabbrielli, V. Bocci, and V. Travagli. 2010. Properties of sesame oil by detailed 1H and 13C NMR assignments before and after ozonation and their correlation with iodine value, peroxide value, and viscosity measurements. Chemistry and Physics of Lipids 163: 148-156.
33. Seo, H. Y., J. Ha, D. B. Shin, S. L. Shim, K. M. No, K. S. Kim, K. B. Lee, and S. B. Han. 2010. Detection of corn oil in adulterated sesame oil by chromatography and carbon isotope analysis. Journal of the American Oil Chemists' Society 87: 621-626.
34. Šmejkalová, D., and A. Piccolo. 2010. High-power gradient diffusion NMR spectroscopy for the rapid assessment of extra-virgin olive oil adulteration. Food Chemistry 118: 153-158.
35. Tay, A., R. Singh, S. Krishnan, and J. Gore. 2002. Authentication of olive oil adulterated with vegetable oils using Fourier transform infrared spectroscopy. LWT-Food Science and Technology 35: 99-103.
36. Torrecilla, J. S., E. Rojo, J. C. Dominguez, and F. Rodríguez. 2010. A novel method to quantify the adulteration of extra virgin olive oil with low-grade olive oils by UV− vis. Journal of Agricultural and Food Chemistry 58: 1679-1684.
37. Varidi, M. J., M. Varidi, M. Vajdi, and A. Sharifpour. 2018. Design, development and application of electronic nose instrument to rapidly detect spoilage of air, vacuum and modified atmosphere packaged camel minced meat. Journal of Food Science and Technology 15.
38. Wojnowski, W., T. Majchrzak, T. Dymerski, J. Gębicki, and J. Namieśnik. 2017. Portable electronic nose based on electrochemical sensors for food quality assessment. Sensors 17: 2715.
39. Xu, L., X. Yu, L. Liu, and R. Zhang. 2016. A novel method for qualitative analysis of edible oil oxidation using an electronic nose. Food Chemistry 202: 229-235.
40. Xu, Y., M. Hassan, F. Kutsanedzie, H. Li, and Q. Chen. 2018. Evaluation of extra-virgin olive oil adulteration using FTIR spectroscopy combined with multivariate algorithms. Quality Assurance and Safety of Crops & Foods 10: 411-421.
41. Zhao, X., D. Dong, W. Zheng, L. Jiao, and Y. Lang. 2015. Discrimination of adulterated sesame oil using mid-infrared spectroscopy and chemometrics. Food Analytical Methods 8: 2308-2314.

Journal of Agricultural Machinery
Vol. 11, No. 2, Fall- Winter 2021, p. 371-383

نشریه ماشین‌های کشاورزی
جلد ۱۱، شماره ۲، نیمسال دوم ۱۴۰۰، ص ۳۸۳-۳۷۱

Estimation of the Best Classification Algorithm and Fraud Detection of Olive Oil by Olfaction Machine

M. R. Zarezadeh[1], M. Aboonajmi[2*], M. Ghasemi Varnamkhasti[3], F. Azarikia[4]

Received: 08-11-2019
Accepted: 15-01-2020

Introduction

Extra Virgin Olive Oil (EVOO) is one of the most common and popular edible oils which is an important part of the Mediterranean diet. It is a rich source of sterol, phenol compounds and vitamins A and E. EVOO has useful effects on human body and significant reduction of cardiovascular diseases due to these benefits, EVOO is expensive so unfortunately adulteration in EVOO by mixing it with other cheap and low cost and low value oils such as canola, sunflower, palm and etc. is very common. Adulteration leads to health and financial losses and sometimes cause serious illness. Olive oil has various quality levels which depend on different factors such as olive cultivar, storage, oil extracting process etc.

Materials and Methods

There are numerous food quality evaluation and adulteration detection approaches which include destructive and non-destructive methods. Control sample (EVOO) was applied from "DANZEH food industry", Lowshan, Gilan Province. For ensuring that control sample is extra virgin, a sample was tested in "Rahpooyan e danesh koolak Lab." Tehran, Iran; according to "Institute of standards and industrial research of Iran" ISIRI number: 4091 and INSO 13126-2. Eight semi-conductor gas sensors "FIS, MQ3, MQ3, MQ4, MQ8, MQ135, MQ136, TGS136, TGS813 AND TGS822" applied in used olfaction machine. In this study there were 6 treatments: 1-Pure EVOO, 2- EVOO with 5% adulteration, 3- EVOO with 10% adulteration, 4- EVOO with 20% adulteration, 5- EVOO with 35% adulteration and 6- EVOO with 50% adulteration. Adulteration created with ordinary frying oil (including sunflower, canola, and maize oils). Each treatment prepared in seven samples and each sample test was repeated seven times. In this study, olfaction machine, a non-destructive, simple and user friendly System applied. As mentioned, the olfaction machine includes eight different sensors, so each test has eight graphs. Four features (1- Sensor output (mV) in start of odor pulse (refer to fig. 3) 2- Sensor output at the end of odor pulse 3- Average of sensor output during odor pulse and 4- Difference of sensor output at the end and start of start of odor pulse); So 32 features extracted and analyzed and finally effective sensors reported.

Results and Discussion

Histogram and box plot of raw data showed that the data are not normal and need some preprocessing operations. Preprocessing facilitates data analyzing and classifying extracted features. After preprocessing, the standard data, divided into two classes: train data (70%) and test data (30%). Data classified with 4 different classifier models which include: K-nearest neighbors, support vector machine, artificial neural network and Ada-boost. Results showed that KNN method, with 89.89% and SVM with 86.52% classified with higher accuracy. Similarly, the confusion matrix showed the reasonable results of classifying operation. Also, three effective sensors in classifying determined TGS2620, MQ5 and MQ4 respectively, and on the other side, sensors such as MQ3 and MQ8 have the minimum effect on classifying so it is possible to remove these sensors from the sensor array without effective impress on results. This may cause decrease in the olfaction machine price and reduce analyzing time.

Conclusions

Due to increasing adulteration in foods, especially in olive oil and its significant effects on people's health and financial losses, a simple, cheap and non-destructive quality evaluation extended. Results showed that the olfaction machine with metal oxide semiconductor (especially including TGS 2620, MQ5 and MQ4 sensors) can use for classification and adulteration detection of extra virgin olive oil. Evaluation of this system's output leads to higher classification accuracy by using KNN and SVM method for olive oil classification and also fraud detection (5% adulteration).

Keywords: Classification, Fraud detection, Olive oil, Olfaction machine, Quality

1- PhD. Candidate, Department of Agro-technology, College of Abouraihan, University of Tehran, Tehran, Iran
2- Associate Professor, Department of Agro-technology, College of Abouraihan, University of Tehran, Tehran, Iran
3- Associate Professor, Department of Mechanical Engineering of Biosystem, Shahrekord University, Shahrekord, Iran
4- Assistant Professor, Department of Food Technology, College of Abouraihan, University of Tehran, Tehran, Iran
(*- Corresponding Author Email: abonajmi@ut.ac.ir)